Assignment # 2:

* **Speech Recognition Systems (as part of MT):**

Speech recognition (also known as automatic speech recognition or computer speech recognition) The term "voice recognition" is sometimes used to refer to speech recognition where the recognition system is trained to a particular speaker - as is the case for most desktop recognition software, hence there is an element of [speaker recognition](http://en.wikipedia.org/wiki/Speaker_recognition), which attempts to identify the person speaking, to better recognize what is being said. Speech recognition is a broad term which means it can recognize almost anybody's speech - such as a call-centre system designed to recognize many voices. Voice recognition is a system trained to a particular user, where it recognizes their speech based on their unique vocal sound. Speech [recognition](http://en.wikipedia.org/wiki/Recognition) applications include voice dialing (e.g., "Call home"), call routing (e.g., "I would like to make a collect call"), [domotic](http://en.wikipedia.org/wiki/Domotic) appliance control and content-based spoken audio search (e.g., find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), preparation of structured documents (e.g., a radiology report), speech-to-text processing (e.g., [word processors](http://en.wikipedia.org/wiki/Word_processor) or [emails](http://en.wikipedia.org/wiki/Email)), and in aircraft [cockpits](http://en.wikipedia.org/wiki/Cockpit).

http://en.wikipedia.org/wiki/Speech\_recognition

* **Voice Recognition Systems :**

Voice recognition is an alternative to typing on a keyboard. Put simply, you talk to the computer and your words appear on the screen. The software has been developed to provide a fast method of writing onto a computer and can help people with a variety of disabilities. It is useful for people with physical disabilities who often find typing difficult, painful or impossible. Voice recognition software can also help those with spelling difficulties, including users with dyslexic, because recognised words are always correctly spelled.

Voice recognition software:

Voice recognition software programs work by analyzing sounds and converting them to text. They also use knowledge of how English is usually spoken to decide what the speaker most probably said. Once correctly set up, the systems should recognise around 95% of what is said if you speak clearly. Several programs are available that provide voice recognition. These systems work best on Windows XP Windows Vista.

A number of voice recognition programs can be used with Windows, including the one supplied with Microsoft Vista. Most specialist voice applications include a software CD, a microphone headset, a manual and a quick reference card.

You connect the microphone to the computer, either into the soundcard (sockets on the back of a computer) or via a USB connection. Then you can begin talking using the following steps.

http://www.macspeech.com/extensions/faq/kb.php?article=20

**Benefits & Areas of Applications:**

-Speech recognition technology seems to have exploded throughout multiple industries.

- Speech recognition technology can also reduce the number of live calls.

- Speech recognition can also reduce the amount of time an agent needs to remain on a line, increasing the number of calls that agent can effectively handle within their shift.

-Voice Recognition Software can make you smarter.

-Voice Recognition Software can make you seem smarter and allow you to impress your friends.

-Voice Recognition Software gives you more free time.

-Voice Recognition Software can improve your social life.

- Voice Recognition Software can cure your medical problems.

- Voice Recognition Software can allow you to surf the Internet hands free.

<http://www.tmcnet.com/channels/speech-recognition/articles/20416-benefits-effective-speech-recognition.htm>

<http://www.voicerecognition.com/talking/benefits.html>

* **Limitations:**

All speech recognition programs (both Macintosh-based and Windows-based) which support transcription from an audio file have the following limitations:

1. **There must be only one voice in the audio file.** To speech recognition software the audio file is just so much data. The program has no way telling when different people are speaking, or switching between voice profiles.
2. **There must be a matching voice profile for the person speaking in the recording.** Without a matching voice profile Listen has no proper references to compare the incoming audio against.
3. **The person speaking must say any necessary punctuation such as commas, periods, question marks, et cetera.** Spoken punctuation is necessary to provide additional context for the words.
4. **It is not a tape recording**; tape hiss causes too much distortion.
5. **You must have the Transcription Pak installed**.

http://www.macspeech.com/extensions/faq/kb.php?article=20

Examples:

* **One example of: text-to-speech online system(&features):**

Voice Forge: Voice Forge makes the Internet Talk. How do we do that? We make it easy for online application developers to add high quality Text-To-Speech (TTS) audio to any Internet site or device. Traditionally, online TTS synthesis has been painful for three reasons:

* **Text-to-Speech software is complex**
* **The voices sound corporate and boring**
* **The audio distribution rights are expensive**

Voice Forge makes online TTS easy through our hosted speech API. You send us text, tell us the voice and we return an MP3 - yes it's that simple. Register as a developer and we'll be happy to share documentation, code examples and expertise on our hosted speech API and service.

Ok, so we put TTS online and made it easy and affordable. But there’s more… the best part is we've created over 45 fun and entertaining voices and we're adding more voice variety each day. Hear all the voice avatars at our online TTS demo page. The whole game changes when **users get to pick a voice they like.**

Why limit yourself with a la carte voices or pay high audio distribution fees? Voice Forge customers get access to all voices for one set price. As long as you are a Voice Forge subscriber you can distribute synthesized audio over the Internet. Pick the plan that works for you and [sign up online](http://www.voiceforge.com/Speech-API.html) today.

<http://www.voiceforge.com/Online-TTS.html>

* **One example of: speech-to-text online system(&features):**Free

**Natural Reader is a Text to Speech software with natural sounding voices. This easy to use software can convert any written text such as MS Word, Webpage, PDF files, and Emails into spoken words. Natural Reader can also convert any written text into audio files such as MP3 or WAV for your CD player or iPod.**  
  
**What Can Natural Reader do for you?**

**1\_** [**Allows students to listen to class notes, text books…etc**](javascript:void(0))

**2\_** [**Facilitates education**](javascript:void(0))

**3\_** [**Learn English or other languages**](javascript:void(0))

**4\_** [**Avoids eyestrain from too much reading**](javascript:void(0))

**5\_** [**Make proofreading effective**](javascript:void(0))

<http://www.naturalreaders.com>

* **Arabic enabled speech recognition software ( clarify: Free or Not):**

Sakhr provides software for Arabic Text-to-Speech (TTS) and Automatic Speech Recognition (ASR). Sakhr TTS converts Arabic text into a natural, human-sounding synthetic voice. Sakhr ASR engine recognizes spoken Arabic in different accents, without requiring user training. Sakhr ASR & TTS solutions include:

* Speech to Speech Translation
* Mobile Applications
* Call Center IVR (Interactive Voice Response)
* Directory Assistance
* Arabic Language Learning

#### Text-to-Speech (TTS)

The Sakhr TTS engine converts Arabic text into human voice. Sakhr TTS is the industry leader in synthesizing a natural, human-sounding Arabic voice. Sakhr TTS leverages the company’s 15+ years research & corpus in [Arabic natural language processing](http://international.sakhr.com/arabic-nlp-natural-language-processing.html) (NLP). This research is critical to overcoming the TTS challenges of the Arabic language, such as lack of diacritics and punctuation. Features:

* Most natural, intelligible Arabic voice in market
* Embedded TTS engine for mobile devices
* Rules-based powerful diacritizer with 97% accuracy
* Normalizes ambiguous text such as dates, time, currencies, abbreviations
* Unlimited vocabulary, text size, raw text, phonetic and prosodic input

#### Speech Recognition (ASR)

Sakhr’s ASR engine is the core technology recognizing spoken Arabic across a wide range of utterances, noisy environments, and Arabic accents from the Arabian Gulf, Egypt and Levant. This engine is industry-standards compliant and can be integrated with mobile client or telephony applications. Features:

* Speaker independent
* High accuracy recognition for large or small vocabularies
* Supports different Arabic accents such as Gulf, Levant & Egyptian
* Integrates with telephony systems for IVR, directory assistance, voice dialing
* Integrates with mobile devices & client applications
* Supports Microsoft Speech Server
* Supports proprietary API and MS SAPI 5.
* Supports open standards such as MS SALT.

<http://international.sakhr.com/arabic-speech-recognition-and-arabic-TTS.html>

* Explore the area of: Speech-to-speech systems=give an example other than what was already handled in the class/ document: